An Energy Efficient Genetic Approach for Clustering of Wireless Sensor Network

Sudakshina Dasgupta[1], Paramartha Dutta[2]

Department of Information Technology[1], Department of Computer and System Science[2]
Government College of Engineering & Textile Technology, Serampore, India[1], Visva Bharati University, Santiniketan, India[2]
[1]sudakshinadasgupta@yahoo.com; [2]paramartha.dutta@gmail.com

***Abstract*-Wireless Sensor Network has emerged as an important technology of the future due to its potential for application across a wide array of domains. The issue of optimizing the limited and often non-renewable energy of sensor nodes due to its direct impact on network lifetime dominates every aspect of wireless sensor networks. Clustering protocols significantly impact network lifetime, routing algorithms and connectivity. In our work we have proposed energy efficient approach for clustering which extends network lifetime while guaranteeing minimum connectivity. Our protocol achieves energy efficiency by reducing the number of Cluster-Head at each round of a sensor network. A genetic algorithm (GA) is used to create energy efficient clusters for data dissemination in wireless sensor networks. The proposed techniques offer better utilization of energy so that the lifetime of the sensor nodes would increase. Here, we are using the method of evolutionary computing for the selection of the Cluster-Heads. The Base Station periodically executes the proposed algorithm to select new Cluster-Heads after a certain period of time.**

Keywords-Wireless Sensor Network; Cluster-Heads; Genetic Algorithm

I. INTRODUCTION

A sensor network is composed of a large number of sensor nodes, which are densely deployed either inside the phenomenon or very close to it. The position of sensor nodes does not need to be engineered or pre-determined. On the other hand, it also means that sensor network protocols and algorithms must possess self-organizing capabilities. Sensor networks may consist of many different types of sensors such as seismic, low sampling rate magnetic, thermal, visual, infrared, acoustic and radar [1]. Sensor nodes can be used for continuous sensing, event detection, event ID, location sensing, and local control of actuators. The concept of micro-sensing and wireless connection of these nodes promise many new application areas, like social, military and environmental problems including secretly monitoring enemy activities on a military battlefield located in an inhospitable terrain, and detecting wild areas in a densely forested areas [3].

A wireless ad-hoc network consists of nodes that move freely and communicate with each other using wireless links. Ad-hoc networks do not use specialized routers for path discovery and traffic routing [5]. In designing the wireless sensor networks there are some constraints, like the small size, low weight, use of energy and the low prices of sensors [2].Among these factors the amount and manner of using energy is very important issue. On the other hand, the decreased use of energy in wireless sensor networks has a direct relation with their increased longevity and this issue is very important either. For increasing the longevity of a network, the load of energy is distributed equally between the nodes of a network [6]. On the other hand, the load distribution is monotonous. The monotonous distribution of energy load in a network will cause the network nodes lose dissipate their energy in a short distance between each other, and the chronological distance between the death of the first node and the last node will become shorter. The relational protocols have an important role in efficiency and longevity of wireless sensor networks. So, designing energy efficient protocols is the formidable necessity for wireless sensor networks. By using these protocols not only the whole consumed energy in a network will decrease, but also the load of consumed energy will be distributed among the network nodes monotonously. So the longevity of the network is expected to increase. Among the available protocols, hierarchical protocols are greatly economized in consuming energy by the network. In these protocols the network is divided into several clusters and in each of these clusters one node will be introduced as a head cluster [8]. The tasks of these cluster heads are together the data send from the nodes of the cluster, mixing the data and forward those data to the sink. In these protocols selecting a node as a cluster head and mixing the data are greatly efficient in increasing the scalability and longevity of a network. Up to now, several clustering protocols like: LEACH [9, 15], TEEN [14], APTEEN [13], etc., have been proposed for the wireless sensor networks. In this article, we suppose the sensor network like static. The sensors are distributed monotonously in an environment and they are far away from sink.

II. RELATED WORK

Heizelman et al. describe the LEACH [15] protocol, as a hierarchical and self-organized cluster-based approach for monitoring applications. The data collection area is randomly divided into several clusters, where the number of clusters is predetermined. Based on time division multiple accesses (TDMA), the sensor nodes transmit data to the cluster heads (CH), then aggregate and transmit the data to the Base Station (BS). A new set of CHs are chosen after specific time intervals. A node can be reelected only after all the remaining candidates have been elected. Lindsey et al. proposes PEGASIS [10], an extension of LEACH, in which nodes transmit to their nearest neighbours and eventually transmit the messages to the base station. Bandyopadhyay and Coyle describe a multilevel hierarchical clustering algorithm [7], in which the parameters for minimum energy consumption are obtained using stochastic geometry. Li Qing, Qingxin Zhu, Mingwen Wang propose DEEC [11], where the cluster heads are elected by a probability based on the ratio between residual energy of each node and the average energy of the network. The approaches of selecting cluster-heads for nodes in a WSN are different according to the initial and residual energy of the nodes. The nodes with high initial and residual energy will have more chances to become the cluster-heads than the nodes with low energy [8, 16].

III. GENETIC ALGORITHM

Genetic algorithms (GAs) [4, 18] are adaptive methods which may be used to solve search and optimization problems. They are based on the genetic processes of biological organisms. According to the principles of natural selection and survival of the fittest, natural populations are evolved in many generations. By mimicking this process, genetic algorithms are able to use solutions to real world computational problems, if they have been suitably encoded. GAs and the population of "individuals", represent a possible solution to a given problem. Each individual assigns a "fitness score" according to how good a solution to the problem is. The highly-fit individuals are given opportunities to "reproduce", by "cross breeding" with other individuals in the population. It produces new individuals as "offspring", which share some features taken from each "parent". The least fit members of the population are less likely to be selected for re-production, so they will "die out". A new population of possible solutions is produced by selecting the best individuals from the current "generation", and by mating them we can produce a new set of potential individuals. This new generation contains a higher proportion of the characteristics possessed by the good members of the previous generation. In this way, in many generations, good characteristics are spread throughout the population. By favouring the mating of the fit individuals, the most promising areas of the search space are explored [17]. If the GA has been designed well, it is expected to converge to an optimal solution to the problem in reasonable amount of time.

A. Application of Genetic Algorithm in Our Proposed Method

The BS uses a GA to create energy-efficient clusters for a given number of transmissions. The node is represented as a bit of a chromosome. The head and member nodes of a cluster are represented by 1s and 0s, respectively. A population consists of several chromosomes. The best chromosome is used to generate the next population. Based on the survival fitness, the population transforms into the future generation. Initially, each fitness parameter is assigned an arbitrary weight. However, after every generation, the fittest chromosome is evaluated and the weights for each fitness parameter are updated accordingly. The GA outcome identifies suitable clusters for the network. The BS broadcasts the complete network details to the sensor nodes. All the sensor nodes receive the packets broadcasted by the BS and clusters are created accordingly. Thus the cluster formation phase will be completed. This is followed by the data transfer phase. The functionality of the system has been described in Fig. 1.

B. Encoding of Our Problem Framework

Finding appropriate cluster heads is critically important for minimizing the distance. We use binary representation in which each bit corresponds to one sensor or node. "1" means that corresponding sensor is a cluster-head; otherwise, it is a regular node. The initial population consists of randomly generated individuals. GA is used to select cluster-heads.

C. Selection

The selection process determines which of the chromosomes from the current population will mate (crossover) to create new chromosomes. These new chromosomes join to the existing population. This combined population will be the basis for the next selection. The individuals (chromosomes) with better fitness values have better chances of selection. There are several selection methods, such as: "Roulette-Wheel" selection, "Rank" selection "Steady state" selection and "Tournament" selection. In Roulette-Wheel, which is used in our work, chromosomes with higher fitness compared to others will be selected for making new offspring with proportionately higher probability.

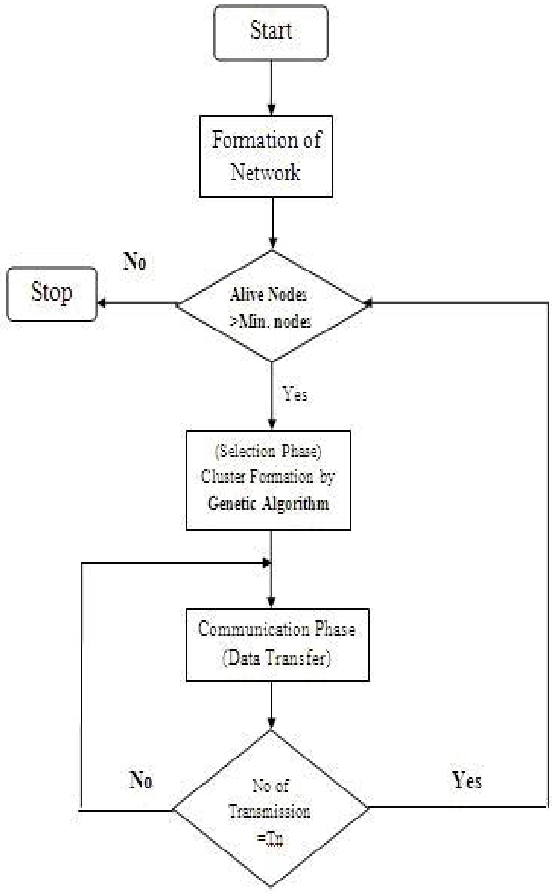

Fig. 1 Flowchart showing the steps of the proposed scheme

D. Crossover and Mutation Used in Our Scheme Crossover

In this work, we have used one-point crossover. If a regular node becomes a cluster-head after crossover, all other regular nodes should check if they are nearer to this new cluster-head. If so, they switch their membership to this new head. This new head is detached from its previous head. If a cluster-head becomes a regular node, all of its members must find new cluster-heads. Every node is either a cluster-head or a member of a cluster-head in the network. An example of crossover operation is shown in Fig. 2.

	First	Second
Chromosome	11001 | 01100000110	10010 | 10110001110
Offspring	11001 | 10110001110	10010 | 01100000110

Fig. 2 An example of Crossover

E. Mutation

The mutation operator is applied to each bit of an individual chromosome with a probability of mutation rate. When applied, a bit whose value is 0 is mutated into 1 and vice versa. The mutation operation is described in Fig. 3.

	Offspring 1	Offspring 2
Original	1100110110001110	1001001100000110
Mutated	1100110010001110	1001001100000110

Fig. 3 An example of Mutation

F. *Fitness Parameter*

The total transmission distance is the main factor in our problem, which needs to be minimized. In addition, the number of cluster heads can factor into the function. Given the same distance, fewer cluster heads result in greater energy efficiency. The fitness function we have calculated in our work is based on the following parameters:

1) *Cluster Distance (C)*:

The cluster distance C is the sum of the distances from the nodes to the cluster head and the distance from the head to the BS. For a cluster with k member nodes, the cluster distance C is defined as follows:

$$C = \sum_{i=1}^{k} d_{ih} + d_{hs}$$

Where dih is the distance from node i to the cluster-head h and dhs is the distance from the cluster-head h to the BS node s.

2) *Transfer Energy (E):*

Transfer energy, E, represents the energy consumed to transfer the aggregated message from the cluster to the BS. For a cluster with k member nodes, cluster transfer energy is defined as follows:

$$E = \sum_{j=1}^{k} E_{T_{jh}} + k \times E_R + E_{T_{hs}}$$

The first term of the above equation shows the energy consumed to transmit messages from k member nodes to the cluster head. The second term shows the energy consumed by the cluster head to receive k messages from the member nodes. Finally, the third term represents the energy needed to transmit from the cluster head to the BS.

3) *Number of Transmissions (T):*

The number of transmissions T is assigned by the BS, for each data transfer stage. The value of T can be adjusted according to the network conditions and current energy levels. Moreover, larger values of T indicate that the outcome of the GA will be used for a longer period of time. In other words, T should be adjusted to reinforce the outcome of a superior solution. The quality of best chromosome is determined from the history of previous GA solutions.

G. *Fitness Function Calculation*

The used energy for conveying the message from cluster to sink and the sending distance are the main factors that we need to minimize them. In addition to these, we can insert the decreasing number of clusters in our function so that it can affect the energy function like the decreased sending distance, because the clusters use more energy in spite of other nodes. So the chromosome fitness or F is a function (Fitness Function) of all the above fitness parameters and can be described as follows:

$$F = (W_1{*}C + W_2{*}E + W_3{*}T)/E_0$$

Here C is the Cluster Distance, E is the Transmission Energy, T is the Number of Transmissions and E_0 is the Initial Energy of the node W_1, W_2 and W_3 are the weight factors for the corresponding system parameters and $W_1+W_2+W_3 = 1$

IV. EXPERIMENTAL RESULTS

The experiments were carried out to measure the performance of the suggested method based on the parameter settings given in Table Ⅰ. We have carried out simulation in MATLAB environment. First, we compare the importance of different parameters in the fitness function. Then, we compare clustering using LEACH algorithm with our proposed method. Fig. 4 shows the simulated environment with number of sensor nodes varying between 100, 500 and 1000 before applying the proposed clustering algorithm. Fig. 5 shows the effect of applying the clustering algorithm to partition the network into number of clusters. In Fig. 6 it can be observed that the network has been partitioned into number of clusters and the number of cluster heads is with in a range of 25 after 2000 generations.

TABLE Ⅰ THE USED PARAMETERS IN OUR EXPERIMENT

Parameter	Value
N	100/500/1000
Population Size	100/200/400
Selection Type	Roulette-Wheel
Crossover Type	One point
Crossover Rate	0.8
Mutation Rate	0.1
Number Of Iteration	1000/ 2000/ 4000

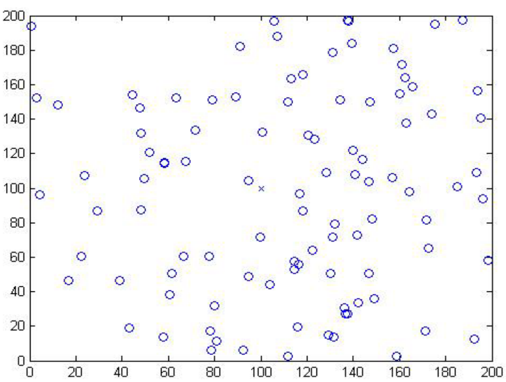

Fig. 4 Nodes deployed at initial phase

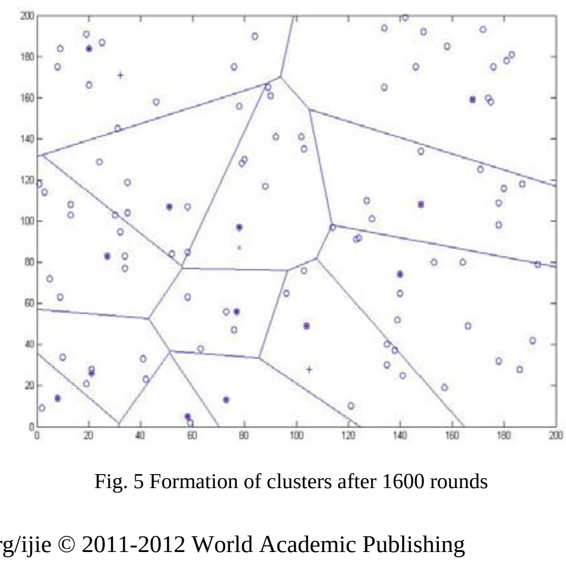

Fig. 5 Formation of clusters after 1600 rounds

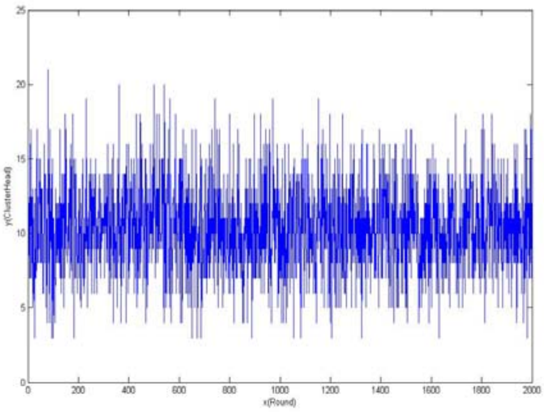

Fig. 6 No. of clusters vs. simulation time after 2000 round

Hence it can be concluded that the proposed method is effective for maintaining a moderate number of cluster heads for a maximum period of time in a cluster based Wireless Sensor Network. Further the performance of the proposed method has been compared with the existing LEACH, M-LEACH and DECC algorithm as shown in Figs. 7, 8 and 9. It can be observed empirically that the proposed method outperforms the existing three methods in terms of number of dead nodes in the network after certain number of rounds, energy reduction rate of the nodes as well as the network after certain number of rounds. Hence it can be concluded that the proposed GA based cluster head selection approach enhances the network lifetime by selecting the cluster heads in a network efficiently and in optimal way as well as rotating the cluster heads in the network as and when required.

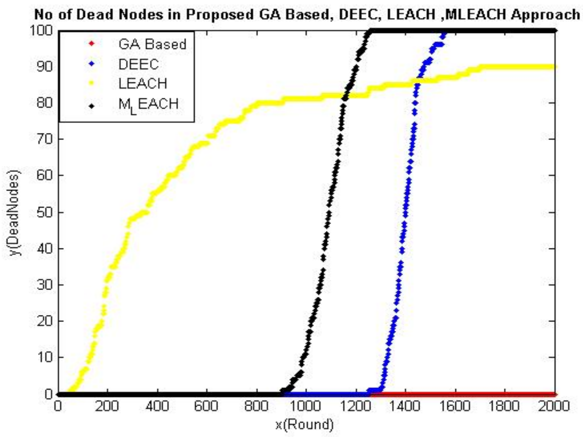

Fig. 7 Number of dead nodes vs. simulation time (round)

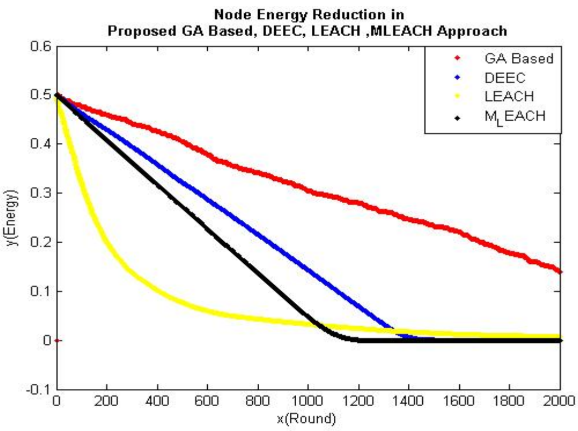

Fig. 8 Decay of node energy vs. simulation time (round)

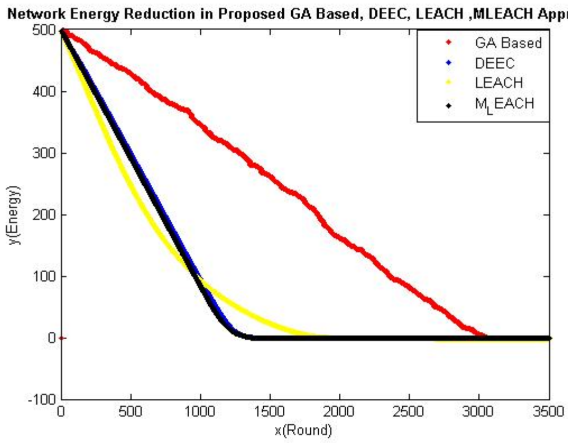

Fig. 9 Decay of network energy vs. simulation time (round)

V. CONCLUSION

In this paper, one method for clustering the network based on the genetic algorithm has been proposed. The basis of this method is the intelligent selection of cluster heads in the network in an optimal way to enhance the network lifetime and also to rotate the cluster heads only when it is required. The results of the experiments in this work demonstrate that the suggested method results in an efficient solution for solving the problem of clustering and cluster head selection in a wireless sensor network, though considerable amount of computation is involved to run the algorithm periodically at the end of the base station.

REFERENCES

[1] I.F.Akyildiz, W.Su, Y.Sankarasubramniam, and E.Cayirei, "A survey on sensor networks, "IEEE Commun Mag., vol.40, no.8, pp.102-114, Aug.2002.

[2] C-Y. Chong, S.P. Kumar., "Sensor Networks: evolution, opportunities and challenges", Proceeding of the IEEE 91(8) (2003)1247-1256.

[3] H. Wang, et al., "Target classification and localization in habitat monitoring ", Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing Hong Kong china., April 2003.

[4] D.E. Goldberg., "Genetic Algorithms in Search, Optimization and Machine Learning", Addison Wesley, New York, 1989.

[5] A.B. McDonald, T. Znati.,"A mobility based framework for adaptive clustering in wireless adhoc networks", IEEE Journal on Selected Areas in Communications 17 (8) (1999) 1466-1487.

[6] G. Gupta, M. Younis., Load-balanced cluster of the International Conference on Communication, Anchorage, Alaska, May 2003.

[7] S. Bandyopadhyay , E. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks", Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies, San Francisco, California, April 2003.

[8] O. Younis, S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed clustering approach for Ad Hoc sensor networks", IEEE Transactions on Mobile Computing 3 (4) (2004)366-379.

[9] W. B. Heinzelman, A. P. Chandrakasan, H. Balakrishnan, "Application specific protocol architecture for wireless microsensor networks", IEEE Transactions on Wireless Networking (2002).

[10] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power-efficient gathering in sensor information systems", Proceedings of the IEEE Aerospace Conference March 2002.

[11] Li Qing, Qingxin Zhu, "Mingwen Wang Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks".

[12] D. Turgut, S. K. Das, R. Elmasri, and B. Turgut, "Optimizing clustering algorithm in mobile adhoc networks using genetic algorithmic approach," Proceedings of the Global Telecommunications Conference (GLOBECOM), November 2002.

[13] Manjeshwar A. and D. P. Agarwal, "APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks", Proc. of the IEEE IPDPS, Fort Lauderdale, USA, Apr. 2002, pp. 195-202.

[14] Manjeshwar A. and D. P. Agarwal, "TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks", Proc. of the IEEE IPDPS, San Francisco, USA, Apr. 2001, pp 23-26.

[15] W.Heizelman, A.Chandrakasan, H.Balakrishman, "Energy-efficient routing protocols for wireless micro sensor networks", Proc33rd Hawaii international conferences on system sciences (HIcS'00)2000.

[16] Jamshid Shanbehzadeh,Saeed Mehrjoo,Abdolhossein Sarrafzadeh ,"An Intelligent Energy Efficient Clustering in Wireless Sensor Networks" ,Proceedings of the International Multiconference of Engineers and computer Scientists 2011 Vol I,IMECS arch 16-18,2011,Hong Kong.

[17] S.Hussain, A.Matin and O.Islam, "Genetic Algorithm for Hierarchical wireless Sensor Network", Journal of Networks, Vol. 2, No. 5, 2007, pp. 87-97.

[18] Jenn-Long Liu and Chinyu V.Ravishankar, "LEACH-GA: Genetic Algorithm –Based Energy-Efficient Adaptive Clustering Protocol for Wireless sensor Networks", International Journal of Machine Learning and Computing, Vol. 1, No. 1, Aprill 2011.

Sudakshina Dasgupta has done her M.Tech in Computer Science & Engineering in the year of 2002 from the University Of Calcutta, Kolkata, India. She is currently acting as Assistant Professor in Department of IT, Govt. College of Engineering & Textile Technology, Serampore, India. She has more than ten years of teaching and more than five years of research experience. Her research area includes Mobile Ad-hoc network, Wireless Sensor and Wireless Mesh network.

Paramartha Dutta is currently Professor in Department of Computer and System Sciences in Visva-Bharati University, Bolpur, Shantiniketan, West Bengal, India. He did his Bachelors and Masters in Statistics from Indian Statistical Institute, Kolkata, India in 1988 and 1990 respectively. Subsequently, He did his Masters in Computer Science in 1993 from Indian Statistical Institute, Kolkata, India. He did his Ph.D. in 2005 from Bengal Engineering and Science University, Shibpore, India. He is a co-author of 6 books and about 100 research publications. He is a Fellow of IETE, OSI, India. He has also the memberships of ACM, ACCS, CSI, IACSIT, IAPR, IEEE, ISCA, and ISTE.